"Mark the Gene": a Method for Nondestructive Introduction of Marker Sequences Inside the Gene Frame of Transgenes[∇]†

Yuki Morono, ¹§ Wataru Kitagawa, ²* Nobutada Kimura, ¹ Naohiro Noda, ¹ Kazunori Nakamura, ¹ and Yoichi Kamagata ^{1,2}

Institute for Biological Resources and Functions, National Institute of Advanced Industrial Science and Technology, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan, and Research Institute of Genome-based Biofactory, National Institute of Advanced Industrial Science and Technology, 2-17-2-1, Tsukisamu-Higashi, Toyohira, Sapporo 062-8517, Japan²

Received 11 January 2007/Accepted 20 May 2007

A specific marking and detection technique is a fundamental requirement for the safer use of genetically modified (GM) organisms. Here we propose a simple and effective method for directly marking functional transgenes in GM organisms. For that purpose, we introduced nucleotide substitutions (NS), based on the degeneracy of codons as markers (NS markers), into the bphC (2,3-dihydroxybiphenyl dioxygenase) and tomA3 (toluene-ortho-monooxygenase) gene frames using a PCR-based method. No change was observed in the enzyme activity of translated proteins, and alignments with homologous genes showed the uniqueness of the NS markers. Furthermore, we constructed tomA3 variations harboring NS markers in different positions. Although the translational products were identical, the constructed variation genes could be distinguished through their marker patterns by multiplex PCR, showing that NS markers could serve as product-specific tags for identifying individual GM organisms. This direct method of marking the functional transgene provides a simple, low-risk, and robust marking method without causing the gene functions to deteriorate.

Genetically modified (GM) microbes (GMMs) have not been widely discussed other than in the heated debate on the safety and regulation of GM plants. Although GMMs have been used for 20 years in the production of pharmaceuticals, vaccines, and enzymes, none of these applications involves the release of GMMs or viruses (31) and only a few GMMs have been approved for release into the environment (8, 27, 32, 35).

Why are GMMs seldom released into the environment despite the relative simplicity and ease of genetic modification? One reason is the small size of the microbes and their diversity, which make them difficult to trace and make it difficult for their ecology to be studied in the environment (31). Specific tracing is the key technology for studying the effectiveness and predicting the fate of GMMs in the environment. Specific tracing is also related to the proper risk management, regulation, and operation of GMMs for safer use in field application (24, 35). The current DNA-based detection method relies mostly on marker sequences, such as antibiotic resistance genes, the transgenes of other functions, or functional sequences (e.g., promoter or terminator) (3). These marker sequences also occur naturally in microorganisms in soil, and some components (e.g., common promoters, such as P_{lac}) are used in multiple GMMs. The presence of a naturally occurring background as well as the same marker sequence in different GMMs can thus cause false-positive results (1). Because of such uncertainty, additional confirmation tests that employ product-specific PCR methods must be carried out.

In a previous report, Marillonnet et al. proposed a method that included an identity tag in GM plants by introducing a silent DNA sequence (300 bp) that coded alphabetical information based on the customized codon table (22). However, as argued by Pauli and Marillonnet et al. (23, 26), there are still some points to be considered, such as the need for extensive verification of safety, the requirement for additional sequencing, and most importantly, gene-crossing problems caused by horizontal gene transfer (HGT), leading to the separation of tag and transgene. The gene-crossing problems are more serious in GMMs, since gene transfer is suspected to be a common event in microbial communities (10, 21). Therefore, the detection of only the GM host or marker sequence adjacent to the transgene would not be more reliable for tracing GMMs and examining their fate in the environment; a method for directly marking and detecting the transgene is an important technology for accurately detecting GMMs.

Here we propose a simple and effective method for directly marking functional transgenes so that they can be specifically traced or detected in the environment without causing the gene functions to deteriorate. For that purpose, we introduced nucleotide substitutions (NS) based on the degeneracy of codons into the transgene as markers (NS markers) through a PCR-based method. The NS-introduced transgene is different from its family genes in the NS region; thus, it can be specifically detected in the environment. The PCR-based introduction and detection of the marker is a method that can be applied universally to any of the genes that have to be detected specifically in the environment, even in the presence of their parental or relative genes.

^{*}Corresponding author. Mailing address: Research Institute of Genome-based Biofactory, National Institute of Advanced Industrial Science and Technology, 2-17-2-1, Tsukisamu-Higashi, Toyohira, Sapporo 062-8517, Japan. Phone: 81-11-857-8517. Fax: 81-11-857-8980. E-mail: w-kitagawa@aist.go.jp.

[§] Present address: Kochi Institute for Core Sample Research, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Monobe B200, Nankoku, Kochi 783-8502, Japan.

 $[\]dagger$ Supplemental material for this article may be found at http://aem.asm.org/.

[▽] Published ahead of print on 25 May 2007.

4916 MORONO ET AL. APPL. ENVIRON. MICROBIOL.

MATERIALS AND METHODS

Bacterial strains, plasmids, and culture conditions. The bacterial strains and plasmids used in this study are listed in Table SA1 in the supplemental material. Standard DNA techniques for the manipulation of plasmids were basically carried out according to standard protocols (28). *Escherichia coli* strains were grown on LB at 37°C; 1 mM ispropyl- β -D-thiogalactopyranoside, 40 μ g/ml 5-bromo-4-chloro-3-indolyl- β -D-galactoside, 100 μ g/ml ampicillin, and 50 μ g/ml kananycin were added to the medium when needed. The DNA fragment extracted from *Burkholderia cepacia* G4 was a kind gift from H. Futamata (Toyohashi University of Technology). The plasmid pHNA2, containing the *bphC* gene, was a kind gift from M. Fukuda (Nagaoka University of Technology).

Introduction and detection of markers in functional genes. The primers used in this study are shown in Table SA2 in the supplemental material. NS markers were introduced into the functional genes by gene splicing with overlap extension (14) (or linking PCR [15]) (see Fig. 2) with some modifications using Pyrobest DNA polymerase according to the manufacturer's instructions. The cycling conditions for linking PCR were as follows: preincubation at 95°C for 3 min; and a total of 30 cycles at 95°C for 30 s, 60°C for 30 s, and 72°C for 1 min/kb of the targeted region. Two or more sets of primers were designed for each marker introduction. Each set of primers contained the marker sequence as an overlapping region and 20 nucleotides (nt) of the priming region complementary to the original gene (see Fig. 2a). In the first step of the PCR, two (or more) fragments that contained a part of the functional gene were amplified so that they had the marker sequence at one end of the products. The resulting PCR fragments were mixed and linked together by a second PCR (linking PCR) using both ends of the primers. The resulting PCR product was digested by an appropriate restriction enzyme and cloned into the plasmids. For the detection of each marker, a common forward primer and marker-specific reverse primers were designed. PCR detection was conducted using AmpliTaq Gold DNA polymerase under the following conditions: preincubation at 95°C for 9 min; and a total of 35 cycles at 95°C for 30 s, 60°C for 30 s, and 72°C for 1 min/kb of targeted region.

Sequence analysis. Nucleotide sequence analyses were performed with a CEQ2000 DNA analysis system (Beckman Coulter). For the sequence data analyses, VectorNTI (Invitrogen, Carlsbad, CA), GENETYX (Software Development, Tokyo, Japan), and GeneWorks (Intelligenetics, Mountain View, CA), as well as the online BLAST homology search (http://www.ncbi.nlm.nih.gov/) and ClustalW multiple sequence alignment (http://www.ddbj.nig.ac.jp/) software packages, were used.

Quantitative PCR. We used a real-time PCR method that uses an internal quenching probe (18, 20) for specific quantification of the marker-introduced gene. The quenching probe was designed to be specific for bphC (5'-GGCCAC GATTGAAAGTAATCTGTATCCTC-3') and labeled with BODIPY FL at its 3' end. The forward and reverse primers for quantitative detection were designed to be specific for NS1 (5'-CGCAGTCGGATGAGAGT-3') and NS2 (5'-GATT CTCTCCATGACATCACC-3'), respectively. All oligonucleotides were obtained from Tsukuba Oligo Service Corp. (Tsukuba, Japan) and used without further purification. PCR with fluorescence monitoring was performed in 20-µl volumes per reaction in a real-time PCR instrument (LightCycler, Roche Diagnostics, Mannheim, Germany), basically following the method reported previously (18). In brief, each reaction mixture contained 1 µM of forward primer, 0.3 μM of reverse primer, 1× TITANIUM Taq PCR buffer, 50 nM of BODIPY FL-labeled internal probe, 200 µM of each deoxynucleoside triphosphate, 0.25 mg/ml bovine serum albumin, and 1 × TITANIUM Taq DNA polymerase (TAKARA BIO INC., Ohtsu, Japan). The standard curve for the quantification was obtained by using the duplicate of the seven-point standards with a known number of target genes (40 to 4.0×10^7). The total community DNAs were extracted from activated sludge (municipal wastewater plant, Kasumigaura, Japan) and lake sediment (Lake Kasumigaura, Japan) using the FastDNA SPIN kit for soil (Obiogene, CA).

Enzyme assays for BphC and TOM. *E. coli* JM109 harboring the recombinant plasmids was grown in LB. The cells were harvested by centrifugation, washed with 50 mM Tris-HCl buffer (pH 7.5), and resuspended in the same buffer. To measure the BphC activity, the cells were disrupted by sonication, and cell debris was removed by centrifugation at 15,000 rpm for 10 min at 4°C. The supernatant was then used as a crude extract. The BphC activity of crude extracts for 2,3-dihydroxybiphenyl (23DHBP) was spectrophotometrically determined by using a DU800 spectrophotometer (Beckman Coulter). Reactions were carried out in 2 ml of 50 mM Tris-HCl buffer (pH 7.5) containing 200 μ M 23DHBP at 25°C. The molecular extinction coefficient of 23DHBP under the assay conditions (pH 7.5) was as follows: $\epsilon_{432} = 13,200 \text{ M}^{-1} \text{ cm}^{-1}$ (2). Protein concentrations were determined with a Bio-Rad Laboratories protein assay kit with bovine serum albumin as the standard. Toluene-*ortho*-monooxygenase (TOM) activity was

measured by naphthol formation from naphthalene by modifying the assay of Canada et al. (4). Two milliliters of washed cells (optical density of 5.0) were contacted with 0.1 mM of naphthalene (dissolved in dimethyl formamide) at 25°C for 30 min. The naphthols synthesized were detected by reacting 800 μl of cell supernatant with 40 μl of 1% (wt/vol) tetrazotized o-dianisidine and 160 μl of glacial acetic acid. The purple azo dye product was detected at an optical density of 528 nm with a DU800 spectrophotometer (Beckman Coulter). The molar amount of naphthols synthesized was calculated by comparison to a 1-naphthol standard curve (molar extinction coefficient of 8,500 M^{-1} cm $^{-1}$ (4). Measurements were done at least three times for each strain tested.

RESULTS

Marker design and its uniqueness. We designed NS markers to be the unique markers in the functional transgene. An NS marker is an assembly of synonymous mutations in a small region (15 to 24 nt) of the functional transgenes. The introduction of the NS marker itself makes no change in the amino acid structure of the translated protein. No preference for the site is needed for NS marker design. With the exception of the tryptophan and methionine codons, at least one base can be substituted. In the codons for serine, arginine, and leucine, 2 or 3 nt can be altered, keeping original coding information unchanged (Fig. 1a).

We targeted the genes bphC, coding for 2,3-dihydroxybiphenyl dioxygenase, and tomA3, coding for the alpha subunit of TOM, as the model target genes. BphC is known to play a key role in the degradation pathway of aromatic compounds in polychlorinated biphenyl-degrading bacteria (17), and TOM is one of the most widely known monooxygenases that is capable of degrading trichloroethylene in Burkholderia cepacia G4 (25). We designed two NS markers for bphC (NS1 and NS2) (Fig. 1b) and six for tomA3 (NS3 to NS8) (Fig. 1c). The alignments of NS1 and NS2 with their corresponding regions in wild-type and related genes showed that the introduction of synonymous mutations differentiated the NS marker sequences from their original or relative sequences (Fig. 1d). In the 15 bases of the marker regions of NS1 and NS2, at least 8 or 5 bases were the mismatches with any of the related sequences. Similarly, at least six bases were the mismatches with any of the related sequences in the marker regions of NS3 to NS8 (see Fig. SA1 in the supplemental material).

Introduction of NS markers and enzymatic activities of the **translated proteins.** The procedure for the introduction of the NS marker was the PCR-based simple method that is similar to the method known as "gene splicing by overlap extension" (14) or "linking PCR" (15). In brief, two sets of PCR primers were designed to have overlapping regions (Fig. 2a) other than the usual priming sites. The substitutions of nucleotides (NS markers) were included in the overlapping region. The first PCR was used to generate two fragments (products AB and CD in Fig. 2b[1] and [2]) that have NS markers at their ends. These two fragments were mixed and used as templates for a second PCR, and the two products were linked to generate an intact marked gene (Fig. 2b[3] and [4]). To our knowledge, at least 25 nt of overlapping region (including 10 nt of additional overlapping regions) was needed for a successful linkage reaction (data not shown).

We obtained the marked genes, namely, $bphC_{\rm NS12}$ and $tomA3_{\rm NS318}$. The gene $bphC_{\rm NS12}$ has NS1 and NS2 markers, and $tomA3_{\rm NS318}$ has NS3, NS4, NS5, NS6, NS7, and NS8 (a total of six) markers. The enzymatic activities corresponding to

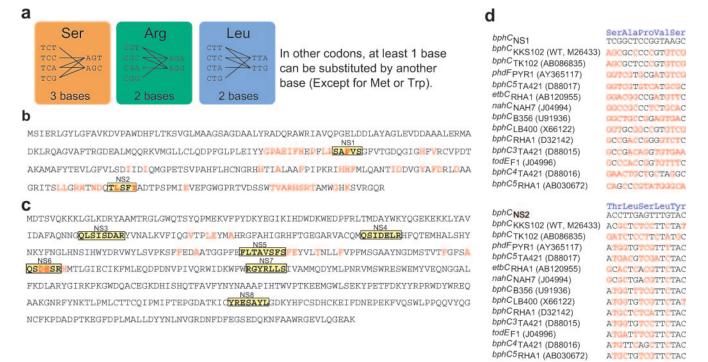


FIG. 1. Marker description and introduction procedure. (a) NS marker that is based on the degeneracy of codons. In Ser, Arg, and Leu, two or three bases can be substituted, and at least 1 base can be substituted in all codons except for Met and Trp. (b and c) The introductory positions of NS markers in the amino acid sequences of BphC and TomA3, respectively. Red letters indicate the essential residues for enzymatic activity or the assembly of subunits derived from information on three-dimensional structures or alignments with known genes for *bphC* and *tomA3*, respectively (19, 29). (d) Sequence alignment of NS1 and NS2 markers with wild-type *bphC* and relative genes. The relative genes were as follows: *bphC*_{KKS102} (wild-type gene in *Achromobacter georgiopolitanum* KKS102, accession no. M26433); *bphC*_{TK102} (*Comamonas testosteroni* TK102, accession no. AB086835); *phhF*_{pyR1} (*Mycobacterium vanbaalenii* strain PYR-1, accession no. AY365117); *bphC*_{5TA421} (*Rhodococcus erythropolis* TA421, accession no. D88017); *etbC*_{RHA1} (*Rhodococcus* sp. strain RHA1, accession no. U91936); *bphC*_{LB400} (*Burkholderia* sp. strain LB400, accession no. X66122); *bphC*_{RHA1} (*Rhodococcus* sp. strain RHA1, accession no. D32142); *bphC3*_{TA421} (*Rhodococcus erythropolis* TA421, accession no. D88015); *todE*_{F1} (*Pseudomonas putida* F1, accession no. J04996); *bphC4*_{TA421} (*Rhodococcus erythropolis* TA421, accession no. D88016); and *bphC5*_{RHA1} (*Rhodococcus* sp. strain RHA1, accession no. AB030672). Red letters indicate the bases that differ from the sequences of the NS markers.

 $bphC_{NS12}$ and $tomA3_{NS3t8}$ were measured and compared with those for wild-type genes. In both of the translated proteins, the activity was almost the same as that of its wild type (104.5 and 100.6% of its wild type activity for BphC and TOM, respectively), showing that NS markers did not cause any deterioration of enzymatic activity (see Table SA3 in the supplemental material). In principle, the NS-marking method introduces a synonymous mutation that does not change the amino acid sequence; thus, the translated proteins are kept intact. In terms of the translation efficiency, the marked genes can be affected by the existence of rare codons in the expression host E. coli (16). In tomA3_{NS3t8}, the introduction of six NS markers resulted in a net increase of 10 rare codons aside from the interchanges between 2 rare codons. However, the negligible change in enzymatic activity of the whole cell indicated that the introduction of the NS markers had a minor effect on translation efficiency in this case.

Specific detection (quantification) by PCR and identification of variations. The NS markers in $bphC_{\rm NS12}$ and $tomA3_{\rm NS318}$ (two and six, respectively) were detected by PCR (Fig. 3a). In the series of PCRs, eight pairs of the primers were used to detect the markers (Fig. 3a, lanes 1 to 8). Among these reactions, the amplified product was observed only in marked

 $bphC_{\rm NS12}$ and $tomA3_{\rm NS318}$ templates. None of the amplified product was detected in the wild-type bphC or tomA3.

In addition, we created six tomA3 variations that have single NS markers (tomA3_{NS3} to tomA3_{NS8}) for "multiplex detection" of several variations (Fig. 3b, lanes 7 to 13). In Fig. 3b, PCRs were performed by using one forward primer and a mixture of the six reverse primers that are specific to the NS3 to NS8 marker sites. The single variations in the reaction tube (single) resulted in single bands (lane 1 to 6). In the case in which a mixture of the several variations was used as the template (multiplexing), multiple bands were observed for the mixture of up to three variations in one reaction. When all six variations were mixed, the intensities of some bands became weak because of the unequal amplification (data not shown). On the other hand, we also created another four variations of marked tomA3 that have multiple markers in one tomA3 gene $(tomA3_{NS34}, tomA3_{NS345}, tomA3_{NS78}, and$ $tomA3_{NS348}$) to make "patterns" of the functional transgene using NS markers like the identification tag (Fig. 3b, lanes 14 to 17). When the tomA3 variations that have a different pattern of the NS marker positioning were used as a template (patterning), multiple bands corresponding to the respective markers were detected, and thus, the marker patterns were successfully read by only a single PCR.

4918 MORONO ET AL. APPL. ENVIRON. MICROBIOL.

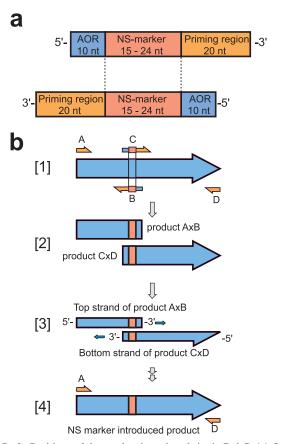


FIG. 2. Positions of the marker-introduced site in BphC. (a) Organization of marker-embedding primer. To our knowledge, at least 25 nt of overlapping region, including 10 nt of additional overlapping regions (AORs), were needed for a successful linkage reaction. (b) Marker introduction procedure for NS marker. The marking was done as follows: [1] two (or more) sets of PCR primers containing marker sequences and an overlapping region for the amplification of gene fragments were designed; [2] partial gene fragments containing marker sequences were generated by PCR; [3] the PCR products were mixed, and denaturation, reannealing, and extension were done to generate hybrid molecules; and [4] PCR amplification with edge primers to obtain the full length of the marker-introduced gene.

We also tried to quantify $bphC_{NS12}$ in the range of 40 to 4 \times 10^6 copies/tube by quantitative PCR in the presence of 4×10^7 copies/tube of wild-type bphC (Table 1). The target sequence (bphC_{NS12}) was successfully quantified in a target/background (wild type) ratio of 0.000001 to 0.1. Moreover, we could quantify a similar range of $bphC_{NS12}$ (400 to 4 × 10⁷ copies/tube) even in the presence of 500 ng of environmental DNA extracted from activated sludge or lake sediment samples. Assuming the microbes to have an average genome size of 4 Mbp, 500 ng DNA corresponds to the genome of 1.1×10^8 microbial cells, and thus, the lower limit for quantification can be calculated as 3 to 4 cells/ml for the case of activated sludge with 10⁹ microbial cells/ml (with an assumption that $bphC_{\rm NS12}$ is on a high-copy-number plasmid [about 100 copies per cell]). In the community experiment, we successfully detected and quantified 10^5 to 10^{10} copies/ml of $bphC_{NS12}$ on the pBBR122 derivative plasmid in Pseudomonas putida KT2440 for more than 30

days in a sequential batch reactor seeded with activated sludge (data not shown).

DISCUSSION

Our method is a direct way to put a mark on the transgene so that it can be specifically traced and precisely identified. Though none of the NS markers has any coded message (22), our marking method has the following advantages. (i) NS markers are introduced inside the gene frame of the transgene, enabling the direct detection of functional transgenes even after gene translocation. (ii) It is possible to create various patterns of transgenes in a marker-dependent fashion while keeping the coding information intact. (iii) If there is any potential risk in this marking method, which is the most important issue for the environmental release of GMMs, it is very low, since the NS marker is composed of only the synonymous mutations in five to eight codons per marker. In addition, the NS markers in the open reading frame of the transgene can be detected using an mRNA-based technique to examine the in situ activity of GMMs. While this method was originally constructed for the marking of the transgene in GMMs, it is applicable for every transgene irrespective of its host organism (plant, mammal, or microorganism).

The NS markers were introduced into the transgenes by the PCR-based method. Many transgenes usually have a reported set of primers that are designed for detection, modification, or other purposes. In our method, two more primers having a marker sequence at each 5' end as the overlapping region are needed for the marker introduction. The marking method itself has no preference in terms of the GM host organism. Since 18 out of 20 amino acids are coded by multiple codons, degeneracy-based codon substitution can thus provide a direct, easy, and flexible marking method for the transgenes. Principally, since the codon substitution is based on synonymous mutations that are thought to be evolutionarily neutral, orthologous genes may have sequences similar to those of the NS markers. However, as shown in Fig. 1d (see also Fig. SA1 in the supplemental material), NS markers showed sufficient specificity among the relative genes.

One of the most attractive points of this marking method is that the NS markers can be introduced while the original amino acid sequence of the transgenes is kept intact. Thus, the introduction of the marker itself can be regarded as self-cloning, so that the marking and detection of particular microbes are possible without the microbes becoming GMMs. The introduction of the NS marker into the transgene causes no change in the activity of the translated protein as long as the translation efficiency is unchanged. As reported by some researchers, rare codons can cause translational problems (6, 12, 16). However, the problem is usually obvious when the same rare codon (e.g., AGA) exists in clusters or in excess numbers in the gene (16). In the $tomA3_{NS3t8}$ gene in this study, the introduction of the 6 NS markers resulted in an increase of 10 rare codons without introducing a cluster of rare codons. The fact that no decrease in enzymatic activity was observed for both BphC_{NS12} and TomA3_{NS3t8} showed that the increase of several rare codons caused no deterioration in gene function. However, the introduction of the NS markers should be considered from the point of view of codon usage in order to

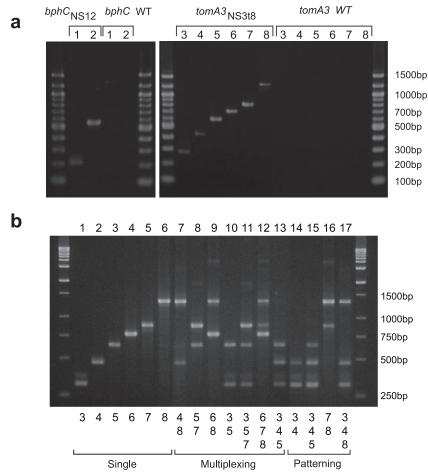


FIG. 3. Qualitative and quantitative detection of the marker sequence by PCR. (a) PCR detection of each of the NS markers with the primers specific for the marker sequence by using each of the marked and wild-type genes as a template. bphC_{NS12}, bphC gene with markers NS1 and NS2; tom43_{NS318}, tom43 gene with markers NS3 to NS8. The numbers on the picture show the primer pair that was used for the detection of each NS marker (with the numbers 1 to 8 corresponding to NS1 to NS8). (b) Identification of the tom43 variations harboring NS markers in different positions. The PCR mixture contained one forward primer and six reverse primers complementary to each of the NS3 to NS8 marker sequences. Lanes 1 to 6, single template with single markers (single detection); lanes 7 to 13, multiple templates with single markers (multiplexing); lanes 14 to 17, single template with multiple markers (patterning). The numbers below the picture show the marker number contained in the respective tom43 variations.

prevent any error of translation, especially in the case of highly expressing proteins.

Another advantage of this method is the capability of creating several variations of particular transgenes that share an identical amino acid sequence. We demonstrated two different uses of such variations, namely, "multiplexing" and "pattern-

TABLE 1. Quantitative detection of $bphC_{NS12}$ by PCR

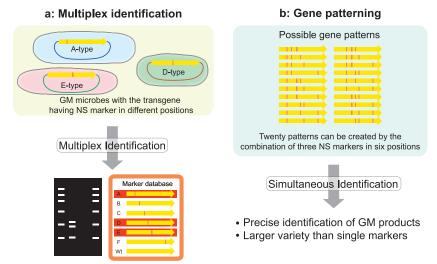
Target/background ratio ^a	Amt of target (no. of copies/tube)	Quantified value ^b
0.000001 0.00001 0.0001 0.001 0.01	4.2×10^{1} 4.2×10^{2} 4.2×10^{3} 4.2×10^{4} 4.2×10^{5} 4.2×10^{6}	$\begin{array}{c} (3.4 \pm 0.95) \times 10^{1} \\ (5.3 \pm 1.30) \times 10^{2} \\ (4.4 \pm 0.73) \times 10^{3} \\ (4.0 \pm 0.60) \times 10^{4} \\ (3.9 \pm 0.44) \times 10^{5} \\ (3.4 \pm 0.26) \times 10^{6} \end{array}$

 $[^]a$ For the quantification, 4.2×10^7 copies of the wild-type bphC gene were added as a background. b Values are means \pm standard deviations of three independent measure-

ing." In multiplexing, six variations of tomA3 harboring single NS markers in different positions were constructed. These six variations could be detected separately by PCR and through multiplex detection of mixtures of up to three variations. Each NS marker was identified by the size of the PCR product. If there are two or more similar GMMs in the environmental samples, the selection of one particular kind of GMM is impractical even just in the mixture of pure cultures. This multiplexed detection system can be a powerful tool for the specific detection and simultaneous identification of the actual producer when there are several different GMMs that share the same transgenes with different NS markers (Fig. 4a). On the other hand, our "patterning" technique can be used for moreprecise identification of GM products. We succeeded in creating several transgenes that differed in terms of their patterns of NS markers while keeping the genetic function intact. To detect the "pattern" of a transgene, only the PCR amplification and identification of the product patterns are needed (Fig. 4b). We succeeded in the PCR detection of up to three different NS markers in one reaction tube. Using combinations of 3 NS

^b Values are means ± standard deviations of three independent measure ments.

4920 MORONO ET AL. APPL. ENVIRON. MICROBIOL.



Simultaneous identification of several types of transgenes

FIG. 4. Schematic diagram of the identification of the GMMs that harbor the same transgene. (a) In multiplex identification, several GMMs with the transgene harboring the NS marker in different positions can be simultaneously detected. (b) In gene patterning, one transgene must be the template DNA. However, 20 patterns of transgenes can be identified only by the patterns of the PCR products.

markers out of 6 positions, 20 variations can be distinguished. The large variety of marker patterns enables the application to a product-specific tag for the transgenes that are widely used in multiple GMMs. The application fields for GM organisms are variable (contained reactor, field release, etc.), and the use of the NS markers can be customized depending on where and how they are used and on the target organism.

Gene dissemination by HGT or gene crossing is a big problem in the tracking of environmental GMMs, and it has been a major issue in the discussion of the impact of GM plants on rhizospheric bacteria (5, 9, 13, 33). There have been several reports on the transfer of the functional gene by homologous recombination (30), and active dissemination of catabolic plasmids by HGT has also been studied as a strategy for bioremediation (7, 34, 36). After the dissemination of the transgene by homologous recombination, the current PCR detection technique, which relies on the primer pair that is complementary to the transgene and its adjacent site (1, 11), is not reliable anymore. However, our NS marker would still be effective even after the genetic crossing, since the introduction sites are inside the gene frame of the functional transgene. The multiple NS markers that can be introduced into a single transgene or all of the transgenes in a particular operon determine the levels for marking depending on their security requirements. In addition, flexible positioning of the NS markers makes it possible to introduce the marker into a structurally important position, such as the substrate-binding site, catalytic center, or elsewhere in a conserved region, allowing more-stable marking. This NS marker is thus a candidate for a flexible, customizable, robust, and widely applicable method for the specific marking of not only GMMs but also all other GM organisms that are to be released in the environment.

ACKNOWLEDGMENTS

This work was supported by the New Energy and Industrial Technology Development Organization (NEDO). W. Kitagawa was supported by a NEDO fellowship.

We thank H. Futamata for providing the DNA fragment extracted from *Burkholderia cepacia* G4 and M. Fukuda for providing the plasmid pHNA2. We also thank O. Yagi and F. Kurisu for helpful advice.

REFERENCES

- Ahmed, F. E. 2002. Detection of genetically modified organisms in foods. Trends Biotechnol. 20:215–223.
- Asturias, J. A., and K. N. Timmis. 1993. Three different 2,3-dihydroxybiphenyl-1,2-dioxygenase genes in the gram-positive polychlorobiphenyl-degrading bacterium *Rhodococcus globerulus* P6. J. Bacteriol. 175:4631–4640.
- Auer, C. A. 2003. Tracking genes from seed to supermarket: techniques and trends. Trends Plant Sci. 8:591–597.
- Canada, K. A., S. Iwashita, H. Shim, and T. K. Wood. 2002. Directed evolution of toluene ortho-monoxygenase for enhanced 1-naphthol synthesis and chlorinated ethene degradation. J. Bacteriol. 184:344–349.
- Castaldini, M., A. Turrini, C. Sbrana, A. Benedetti, M. Marchionni, S. Mocali, A. Fabiani, S. Li, F. Santomassimo, B. Pietrangeli, M. P. Nuti, N. Miclaus, and M. Giovannetti. 2005. Impact of Bt corn on rhizospheric and soil eubacterial communities and on beneficial mycorrhizal symbiosis in experimental microcosms. Appl. Environ. Microbiol. 71:6719–6729.
- Chen, G. T., and M. Inouye. 1994. Role of the AGA/AGG codons, the rarest codons in global gene expression in Escherichia coli. Genes Dev. 8:2641– 2652
- Dejonghe, W., J. Goris, S. E. Fantroussi, M. Hofte, P. De Vos, W. Verstraete, and E. M. Top. 2000. Effect of dissemination of 2,4-dichlorophenoxyacetic acid (2,4-D) degradation plasmids on 2,4-D degradation and on bacterial community structure in two different soil horizons. Appl. Environ. Microbiol. 66:3297–3304.
- De Leij, F., E. J. Sutton, J. M. Whipps, J. S. Fenlon, and J. M. Lynch. 1995.
 Field release of a genetically modified Pseudomonas fluorescens on wheat: establishment, survival and dissemination. Nat. Biotechnol. 13:1488–1492.
- Dunfield, K. E., and J. J. Germida. 2004. Impact of genetically modified crops on soil- and plant-associated microbial communities. J. Environ. Qual. 33:806–815.
- Frost, L. S. 1992. Bacterial conjugation: everybody's doin' it. Can. J. Microbiol. 38:1091–1096.
- Garcia-Canas, V., A. Cifuentes, and R. Gonzalez. 2004. Detection of genetically modified organisms in foods by DNA amplification techniques. Crit. Rev. Food Sci. Nutr. 44:425–436.
- Gurvich, O. L., P. V. Baranov, R. F. Gestel, and J. F. Atkins. 2005. Expression levels influence ribosomal frameshifting at the tandem rare arginine codons AGG_AGG and AGA_AGA in *Escherichia coli*. J. Bacteriol. 187: 4023–4032.
- Heinemann, J. A., and T. Traavik. 2004. Problems in monitoring horizontal gene transfer in field trials of transgenic plants. Nat. Biotechnol. 22:1105– 1109.
- Horton, R. M., Z. L. Cai, S. N. Ho, and L. R. Pease. 1990. Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. BioTechniques 8:528–535.
- 15. Horton, R. M., H. D. Hunt, S. N. Ho, J. K. Pullen, and L. R. Pease. 1989.

- Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene 77:61–68.
- Kane, J. F. 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. Curr. Opin. Biotechnol. 6:494–500.
- Kimbara, K., T. Hashimoto, M. Fukuda, T. Koana, M. Takagi, M. Oishi, and K. Yano. 1989. Cloning and sequencing of two tandem genes involved in degradation of 2,3-dihydroxybiphenyl to benzoic acid in the polychlorinated biphenyl-degrading soil bacterium *Pseudomonas* sp. strain KKS102. J. Bacteriol. 171:2740–2747.
- Kurata, S., T. Kanagawa, K. Yamada, M. Torimura, T. Yokomaku, Y. Kamagata, and R. Kurane. 2001. Fluorescent quenching-based quantitative detection of specific DNA/RNA using a BODIPY((R)) FL-labeled probe or primer. Nucleic Acids Res. 29:E34.
- Leahy, J. G., P. J. Batchelor, and S. M. Morcomb. 2003. Evolution of the soluble diiron monooxygenases. FEMS Microbiol. Rev. 27:449–479.
- Lee, T. H., S. Kurata, C. H. Nakatsu, and Y. Kamagata. 2005. Molecular analysis of bacterial community based on 16S rDNA and functional genes in activated sludge enriched with 2,4-dichlorophenoxyacetic acid (2,4-D) under different cultural conditions. Microb. Ecol. 49:151–162.
- Lorenz, M. G., and W. Wackernagel. 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiol. Rev. 58:563–602.
- Marillonnet, S., V. Klimyuk, and Y. Gleba. 2003. Encoding technical information in GM organisms. Nat. Biotechnol. 21;224–226.
- Marillonnet, S., V. Klimyuk, and Y. Gleba. 2003. Identity tags revisited. Nat. Biotechnol. 21:974–975.
- Morrissey, J. P., U. F. Walsh, A. O'Donnell, Y. Moenne-Loccoz, and F. O'Gara. 2002. Exploitation of genetically modified inoculants for industrial ecology applications. Antonie Leeuwenhoek 81:599–606.
- Nelson, M. J., S. O. Montgomery, W. R. Mahaffey, and P. H. Pritchard. 1987. Biodegradation of trichloroethylene and involvement of an aromatic biodegradative pathway. Appl. Environ. Microbiol. 53:949–954.
- Pauli, U. 2003. Identity tags. Nat. Biotechnol. 21:598.
- 27. Ripp, S., D. E. Nivens, Y. Ahn, C. Werner, I. J. Jarrell, J. P. Easter, C. D.

- **Cox, R. S. Burlage, and G. S. Sayler.** 2000. Controlled field release of a bioluminescent genetically engineered microorganism for bioremediation process monitoring and control. Environ. Sci. Technol. **34**:846–853.
- Sambrook, J., and D. W. Russell. 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Senda, T., K. Sugiyama, H. Narita, T. Yamamoto, K. Kimbara, M. Fukuda, M. Sato, K. Yano, and Y. Mitsui. 1996. Three-dimensional structures of free form and two substrate complexes of an extradiol ring-cleavage type dioxygenase, the BphC enzyme from Pseudomonas sp. strain KKS102. J. Mol. Biol. 255:735–752.
- Springael, D., and E. M. Top. 2004. Horizontal gene transfer and microbial adaptation to xenobiotics: new types of mobile genetic elements and lessons from ecological studies. Trends Microbiol. 12:53–58.
- 31. Stanley, J., K. Hutchinson, I. Godwin, P. Gregg, R. Jessop, M. Katz, K. King, J. Prior, B. Sindel, and D. Thompson. 2006. Biotechnology (GMO) issues and research priorities in natural resource management. Land & Water Australia, Canberra, Australia.
- 32. Thompson, I. P., A. K. Lilley, R. J. Ellis, P. A. Bramwell, and M. J. Bailey. 1995. Survival, colonization and dispersal of genetically modified *Pseudomonas fluorescens* SBW25 in the phytosphere of field grown sugar beet. Nat. Biotechnol. 13:1493–1497.
- Thomson, J. A. 2001. Horizontal transfer of DNA from GM crops to bacteria and to mammalian cells. J. Food Sci. 66:188–193.
- 34. Top, E. M., P. Van Daele, N. De Saeyer, and L. J. Forney. 1998. Enhancement of 2,4-dichlorophenoxyacetic acid (2,4-D) degradation in soil by dissemination of catabolic plasmids. Antonie Leeuwenhoek 73:87–94.
- Urgun-Demirtas, M., B. Stark, and K. Pagilla. 2006. Use of genetically engineered microorganisms (GEMs) for the bioremediation of contaminants. Crit. Rev. Biotechnol. 26:145–164.
- Weekers, F., C. Rodriguez, P. Jacques, M. Mergeay, and P. Thonart. 2001.
 Dissemination of catabolic plasmids among desiccation-tolerant bacteria in soil microcosms. Appl. Biochem. Biotechnol. 91–93:219–232.